

EVALUATION OF DEEP LEARNING TECHNIQUES FOR DEFORESTATION DETECTION IN THE AMAZON FOREST

M. X. Ortega¹, J. D. Bermudez¹, P. N. Happ¹, A. Gomes², R. Q. Feitosa^{1,3},

¹ Pontifical Catholic University of Rio de Janeiro, Brazil - (mortega, bermudez, patrick, raul)@ele.puc-rio.br

² National Institute for Space Research (INPE), São José dos Campos, Brazil - alessandra.gomes@inpe.br

³ Rio de Janeiro State University, Rio de Janeiro-RJ, Brazil

ICWG II/III: Pattern Analysis in Remote Sensing

KEY WORDS: Deep learning, Convolutional Neural Networks, Image Stacking, Deforestation, Amazon Rainforest.

ABSTRACT:

Deforestation is one of the main causes of biodiversity reduction, climate change among other destructive phenomena. Thus, early detection of deforestation processes is of paramount importance. Motivated by this scenario, this work presents an evaluation of methods for automatic deforestation detection, specifically Early Fusion (EF) Convolutional Network, Siamese Convolutional Network (S-CNN) and the well-known Support Vector Machine (SVM), taken as the baseline. These methods were evaluated in a region of the Brazilian Legal Amazon (BLA). Two Landsat 8 images acquired in 2016 and 2017 were used in our experiments. The impact of training set size was also investigated. The Deep Learning-based approaches clearly outperformed the SVM baseline in our approaches, both in terms of F1-score and Overall Accuracy, with a superiority of S-CNN over EF.

1. INTRODUCTION

The Amazon Rainforest accommodates a large biodiversity. It is home to a large number of species, including endemic and endangered flora and fauna. It contains 20% of the fresh water of the planet (Assunção, Rocha, 2019) and produces more than 20% of the world oxygen (Butler, 2008). Therefore, Amazon provides essential resources for the maintenance of our planet (De Souza et al., 2013), (De Souza, Junior, 2018) and its preservation is of paramount importance.

For many years, the Amazon region has faced several threats as a result of unsustainable economic development, such as the extension of agricultural activities at industrial scale (e.g., soybeans, cattle), slash-and-burn land grabbing by underprivileged rural communities, forest fires, illegal gold mining and logging, expansion of informal settlements, and infrastructure construction (roads and train tracks) (Goodman et al., 2019), (Malingreau et al., 2012), (Barreto et al., 2006). Therefore, it is imperative to promote sustainable development to achieve an ecological balance and to contribute to the mitigation of climate change (Sathler et al., 2018). Controlling and monitoring this ecosystem is fundamental to enforce public policies and to avoid illegal activities in the region. Remote sensing has proven to be a cost-effective information source to attain such objectives.

Given the dynamics and complexity of the Amazon region, there have been large government investments aimed at controlling, preventing and combating illegal deforestation (Diniz et al., 2015). The Brazilian National Institute for Space Research (INPE) has developed and maintained a number of projects to provide surveillance reports over the Brazilian Legal Amazon (BLA). The best known-action is the Amazon Deforestation Monitoring Project (PRODES) (Valeriano et al., 2004), which supervises the deforestation in areas with native vegetation of BLA since 1988. The near real-time deforestation detection

(DETER) (Shimabukuro et al., 2007) project, was developed to support land use policies in BLA and controls the illegal deforestation and forest degradation. The Brazilian Amazon Forest Degradation Project (DEGRAD) (Shimabukuro et al., 2015) measures areas in the process of deforestation where the forest cover has not yet been completely removed. Finally, the Land Use and Land Cover Mapping of Amazon Deforested Areas (TerraClass) project (De Almeida et al., 2016) is responsible for qualifying deforestation in BLA and investigating the possible causes of logging. These projects, however, adopt methodologies that involve a lot of manual operations. There is, therefore, a demand for automatic procedures that can improve accuracy and alleviate the human work process, as well as reduce the time needed to generate results.

Numerous change detection techniques have been proposed thus far. Some of the traditional no supervised methods are based on image algebra such as Image Differencing (Jensen, Toll, 1982), Image Ratioing (Howarth, Wickware, 1981), Regression Analysis (Ludeke et al., 1990) and Change Vector Analysis (CVA) (Nackaerts et al., 2005). In addition, techniques based on transformations such as Principal Component Analysis (PCA) (Deng et al., 2008) and Tasseled cap (KT) (Han et al., 2007) have been also used for this purpose. However, these methods require the selection of a proper threshold to identify the changed regions and the features adopted by these conventional algorithms are hand-crafted, which may lead to poor image representations (Zhan et al., 2017).

Support Vector Machine (SVM) is one of the most popular supervised algorithms used in satellite image classification (Dhingra, Kumar, 2019), (Kranjčić et al., 2019) due to its good performance and robustness when labeled samples are scarce. Additionally, random forest (Pal, 2005) and methods based on artificial neural networks (ANN) are also widely used (Maxwell et al., 2018). Recently, Deep

Learning (DL) techniques have been successfully applied to Remote Sensing (RS) image analysis. Using Deep Neural Networks (DNNs), it is possible to learn multiple levels of data representation and to extract more robust and abstract features (Zhan et al., 2017), which usually provide more meaningful information than hand-crafted ones. In this sense, DNNs variants, such as Convolutional Neural Networks (CNNs) and Siamese Networks, are potential candidates for automatic deforestation detection.

In (Zagoruyko, Komodakis, 2015), the authors proposed and explored different CNN architectures to learn similarity functions between images pairs that implicitly suffered some transformations and other kinds of effects (due to e.g., rotation, translation, illumination, etc.). These algorithms presented good performances in comparison to methods based on hand-crafted feature descriptors. Examples of such algorithms are the Early Fusion and the Siamese CNN approaches, which were also used by (Daudt et al., 2018) to detect changes in urban areas. Similarly, (Zhang et al., 2018) successfully applied a Siamese CNN to identify building and tree changes, and also to distinguish between real changes from false ones caused by misregistration errors or false matches.

Moved by the success of DL methods for change detection applications, in this work, we adapt and evaluate Early Fusion and Siamese networks for deforestation detection in the Amazon rainforest. We take as baseline a binary SVM classifier for comparison purposes.

The remainder of this paper is organized as follows. Section 2 presents the change detection methods considered in this work. Section 3 describes the dataset and the adopted experimental protocols. The experimental results are presented in section 4 and some concluding remarks, which also point to future works are included in section 5.

2. CHANGE DETECTION METHODS

In this section, we shortly describe the methods evaluated in this work for deforestation detection: Early Fusion (EF) and Siamese Convolutional Network (S-CNN).

2.1 Early Fusion (EF)

The EF method is inspired by the CNN model proposed in (Daudt et al., 2018), which demonstrated good performance for change detection in urban areas. It is composed of several convolutions and pooling layers, followed by a fully connected (FC) layer, and a softmax layer to carry out the final classification.

The name Early Fusion is related to the concatenation of the images from two different dates, before applying the CNN model. The images are stacked along their spectral dimension to generate a unique input image for patch extraction. These patches are extracted in a sliding windows procedure. Then, the class label is assigned to the central pixel of each patch. The procedure is illustrated in (Figure 1).

2.2 Siamese Network (S-CNN)

The Siamese CNN is an adaptation of a traditional CNN, which comprises two identical branches that share the same hyperparameters and weights values (Zhang et al., 2018).

The architecture adopted in this work is inspired by (Daudt et al., 2018), which was also used for urban changing detection. Both input images are treated independently. Each branch of the Siamese network receives as input one patch cropped from coregistered image pair. The two outputs are concatenated producing the final feature vector (Zhang et al., 2018), (Zagoruyko, Komodakis, 2015). Such vector is the input to a classifier that assigns it to a class: deforestation and no-deforestation. Similar to EF, the class label is assigned to the central pixel of each patch. This process is summarized in Figure 2.

3. EXPERIMENTS

3.1 Data Set Description

The study area is located in BLA, more specifically in Pará State, Brazil, centered on coordinates of 03° 17' 23" S and 050° 55' 08" W. This area has facing a significant deforestation process that has been tracked and monitored by PRODES (Valeriano et al., 2004). Figure 3(c) shows the reference change map of deforestation occurred between December 2016 and December 2017. This data is freely available at the PRODES database (<http://terrabrasilis.dpi.inpe.br/map/deforestation>). However, some polygons of the reference were unconsidered because they had been deforested in the previous years.

The dataset comprises a pair of Landsat 8-OLI images, with 30m spatial resolution. We applied an atmospheric correction to each scene, and then, clipped them to the target area. The final images have 1100 × 2600 pixels and seven spectral bands (Coastal/Aerosol, Blue, Green, Red, NIR, SWIR-1, and SWIR-2). The first image is from August 2nd, 2016 (Figure 3(a)) and the second one from July 20th, 2017 (Figure 3(b)). These dates were chosen due to the lower presence of clouds, a common problem over all BLA region.

3.2 Experimental Setup

Our experiments relied on a pair of optical images acquired approximately one year apart from each other.

In addition, the *Normalized Difference Vegetation Index (NDVI)* was calculated for every pixel as in Equation 1. This index quantifies the presence and quality of vegetation and it is calculated using bands 5 and 4 for Landsat 8, corresponding to the spectral reflectance measurements acquired in the near-infrared and red regions.

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

The *NDVI* was stacked along the spectral dimension of the corresponding images, resulting in images with eight bands. The spectral bands of each image were normalized to zero mean and unit variance. The input to EF was a tensor of a size of 15-by-15-by-16 and to S-CNN a tensor of a size of 15-by-15-by-8 in each branch and the input. We used as baseline a SVM classifier, whose input was a vector of dimension 15 × 15 × 16. In all cases, the patches were extracted using a sliding window procedure with stride equal to three. The window size for each method and the stride size were chosen empirically.

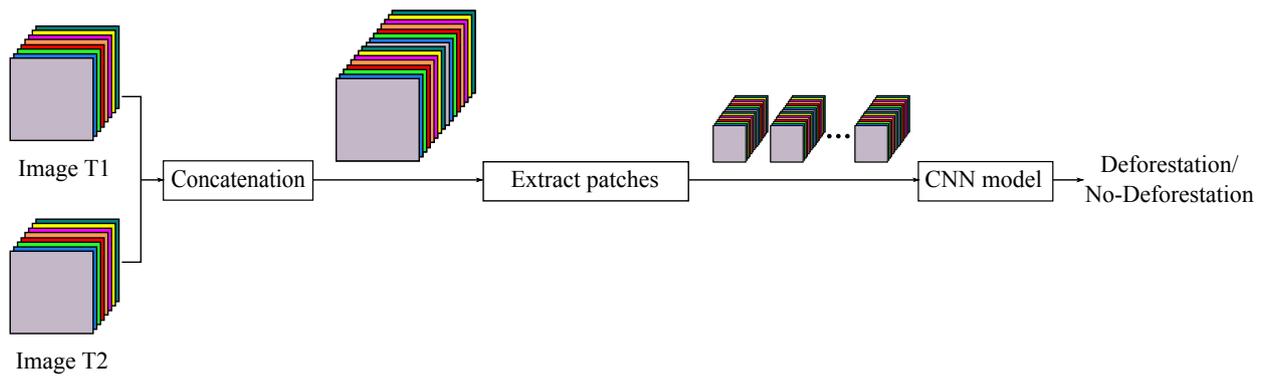


Figure 1. EF approach. Images at different dates (T1 and T2) are concatenated to produce an image pair; then, patches are extracted and fed to the CNN model.

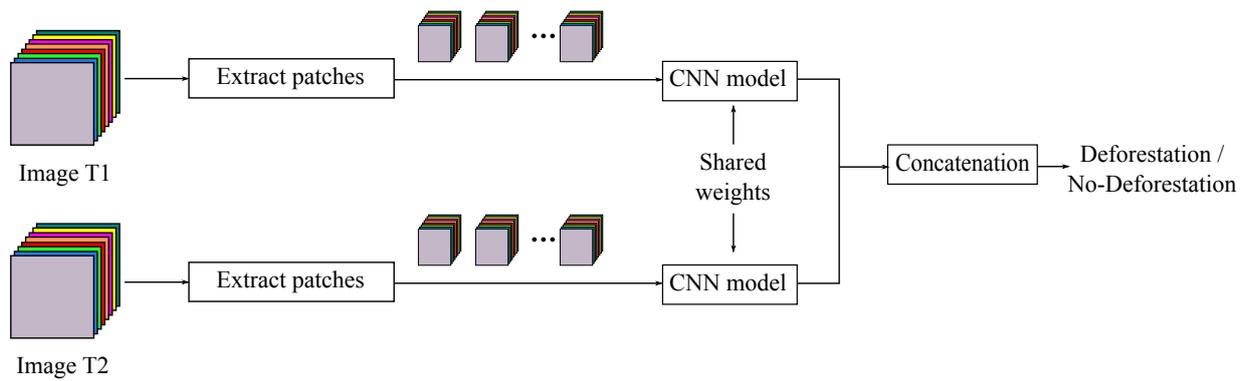


Figure 2. Siamese network. Patches of each image (T1 and T2) are extracted and fed to the CNN model independently. The two branches in the network share exactly the same architecture and parameter values.

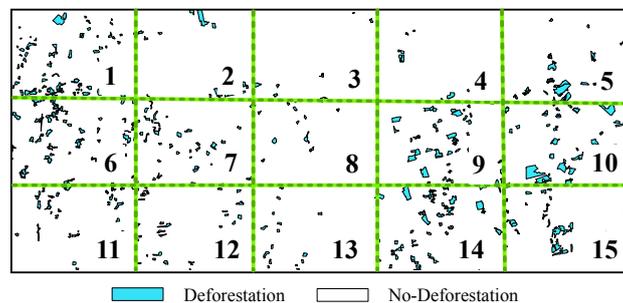
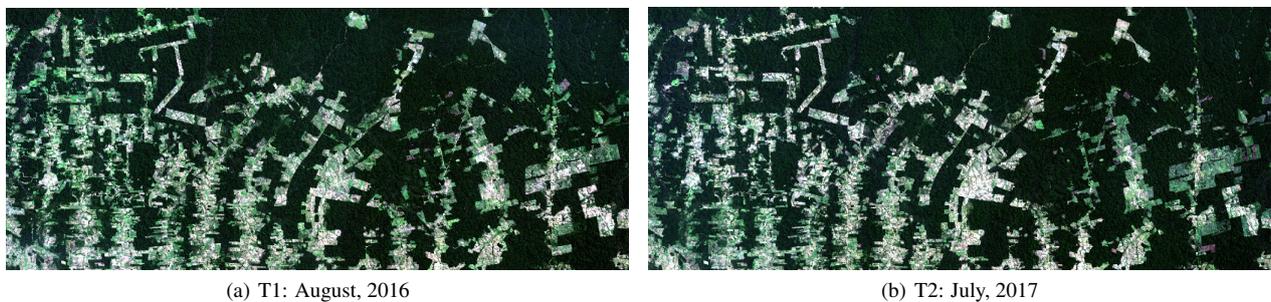


Figure 3. RGB composition of the selected Amazon Forest region at dates T1 (a) and T2 (b); and the deforestation reference set (c). The study area is divided into 15 tiles.

Similar to (Zhang et al., 2018), we divided the input images into tiles. We obtained 15 tiles as shown in Figures 3(a) and 3(b). Tiles 1, 7, 9 and 13 were used for training, tiles 5 and 12 for

validation, and tiles 2, 3, 4, 6, 8, 10, 11, 14 and 15 for testing. The number of available samples of class no-deforestation was

much higher than that of class deforestation. So, we performed data augmentation on samples of deforestation class. Each training pair was rotated by 90° , flipped in the horizontal and vertical axis. In addition, we applied an under-sampling technique on the majority class (no-deforestation) to balance the number of training pairs for both classes. This way, we obtained 8,118 training pairs for each class. The validation set had 40,642 pairs, 963 of the deforestation class and 39,679 of no-deforestation class. This corresponded to the class distribution in the test set, which comprised 1.716,000 pairs, of which 40,392 were deforestation pairs and 1.675,608 no-deforestation pairs.

To assess the influence of the number of training samples, we also considered three different scenarios: using only training samples from a single tile (13), from two tiles (1, 13) and from three tiles (1, 7, 13), yielding 717, 2,127 and 5,421 samples per class, respectively.

We selected the *Radial Basis Function* (RBF) as SVM kernel with the γ parameter set to 0.00027 based on following relationships: $\gamma = \frac{1}{d}$, being d the number of features, as proposed in (Gola et al., 2019). The parameter C was set to 10. This choice was based on a k -fold cross-validation procedure, where k was set to five. The experiments were implemented and carried out in the Python environment using the SVM implementation of the Scikit-Learn (Pedregosa et al., 2011) library.

The CNN architecture used for EF approach is illustrated in Figure 4. It was composed of three Convolutional layers (Conv) with ReLU as activation function, two Max-pooling (MaxPool) layers and two Fully Connected layers (FC) at the end, where the last one is a softmax with two outputs, one associated to deforestation and the other one to no-deforestation class. Regarding the S-CNN model, the two branches comprises the same network architecture Figure 4, but in this case, the network has only a fully connected layer at the end, then, the vectors at the output of each CNN branch were concatenated to compose a new feature vector, which represented the image pair.

The parameter setup of the CNN was: batch size was set to 32 and the number of epochs was set to 100. To avoid over-fitting, we used early stopping to break after 10 epochs without improvement and dropout with rate set to 0.2 in the last fully connected layer. In contrast to (Daudt et al., 2018), where Average Stochastic Gradient Descent (ASGD) was used, we employed the Adam optimizer, which presented a better performance in our preliminary experiments with learning rate of 10^{-3} and weight decay of 0.9.

4. RESULTS

Figure 5 summarizes the results of our experiments in terms of F1-score of class deforestation achieved by the three methods described in Section 2. The figure shows the performance obtained by each method for different number of tiles used for training.

S-CNN achieved the best performance in terms of F1-score in all experiments. As expected, the methods improved their performance as the number of training samples increased. When just one tile was used for training, we recorded F1-scores equal to 46%, 44% and 48%, for SVM, EF and S-CNN respectively. SVM outperformed EF but was still below

S-CNN. This was not unexpected because SVM tends to generalize well under scarce labeled data. In contrast, when two, three and four tiles were taken for training, the EF and S-CNN presented better performance than SVM. With four training tiles, EF and S-CNN outperformed SVM in 10% and 13%, respectively, in terms of F1-Score. Clearly, the DL methods benefited from the increase of training samples than SVM.

We should bear in mind that, in the target application, the classes are highly unbalanced with a predominance of no-deforestation class. Then, under these conditions, the F1-score often tends to decrease for deforestation class.

The results in terms of Overall Accuracy (OA) are presented in Figure 6. As in the F1-Score, the results were improved when the number of training tiles increased. In all scenarios, scores above 90% were achieved. The scores went from about 95%, for one tile, to 97% when four tiles were used. In comparison to the F1-score results, the higher values for OA are related to the higher number of no-deforestation samples that were correctly classified.

Figures 7, 8 and 9 show the RGB composition of tile 2, 6 and 14, respectively, using four tiles for training. They show the tiles in both dates as well as the change maps delivered by each method. The maps show that S-CNN better identified deforested areas (Figures 7-e, 8-e, 9-e). It achieved the highest true deforestation rate, so it presented a lower false deforestation rate than SVM and EF, demonstrating a more accurate result in these three tiles. On the other hand, EF produced the lowest number of false detections, but it did not correctly identify many areas that suffered deforestation as revealed in Figures 7-d, 8-d, 9-d. Notably, much of the false deforestation (reddish) and false no-deforestation (blueish) occurred at the borders of true detected deforested areas (yellowish). This type of error might have resulted from inaccuracies in the delimitation of deforestation polygons. Figures 7-c, 8-c, 9-c shows that SVM presented low performance for deforestation detection. The false deforestation rate was relatively high in all the cases: many pixels were incorrectly identified as deforested areas. We can also observe a salt-and-pepper effect in the SVM outcomes. The same aforementioned trends are presented in the rest of the test area.

5. CONCLUSIONS

This work reported an evaluation of recently proposed deep learning based methods for detection of deforestation in the Amazon forest. Three methods were tested: Early Fusion (EF), Siamese Convolutional Neural Network (S-CNN) and the Support Vector Machine (SVM), the last one taken as the baseline. We used as database a region of the Brazilian Legal Amazon, which has suffered under intense attacks in the last few years. In our experiments, S-CNN was consistently superior to its counterparts in terms of F1-score and Overall Accuracy. The difference to the second approach, EF, was in the range of 3% in terms of F1-score. Actually, in just one experimental setup SVM outperformed EF but not S-CNN by a small margin. Yet, in this case, as in all other tested configurations, S-CNN and EF were much superior to SVM in terms of F1-score.

It is worth mentioning that the performance recorded in our experiments was generally below what was reported in the

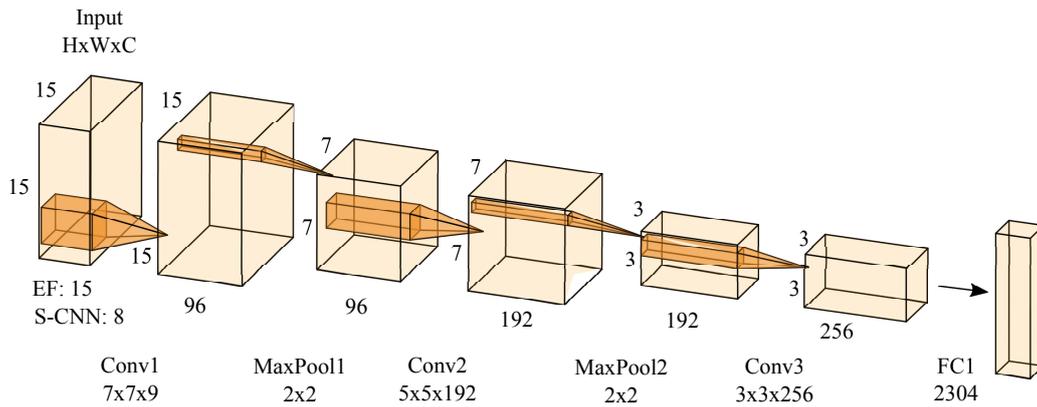


Figure 4. Parameters of the EF and S-CNN architecture.

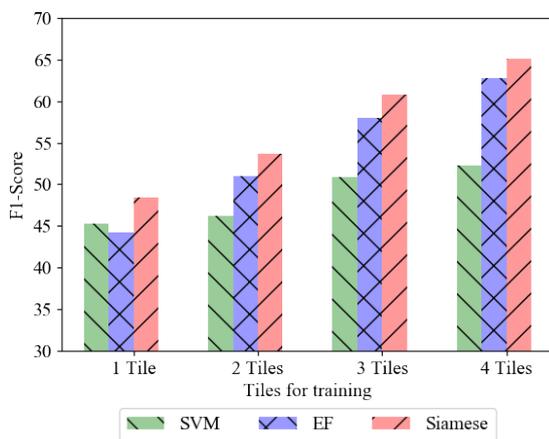


Figure 5. F1-Score from each method using 1, 2, 3 and 4 Tiles for training.

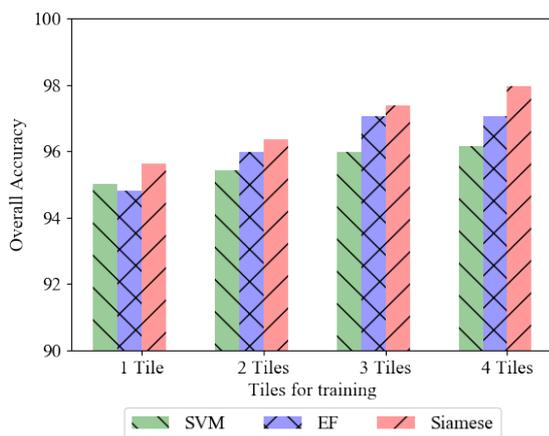


Figure 6. Overall Accuracy from each method using 1, 2, 3 and 4 Tiles for training.

literature for the same methods to detect changes in urban areas. This indicates that deforestation detection is comparatively a more challenging task and more research is required to obtain viable operational automatic solutions.

Future works are intended to fine-tune the hyperparameters of

the tested methods in order to reduce the false deforestation rate. Another investigation is related to the usage of a sequence of images for change detection instead of only an image pair, as well as the usage of Synthetic Aperture Radar (SAR) data. Indeed, the Amazon region is covered by clouds most of the year, which limits considerably the usage of optical data. Under these conditions, SAR data becomes an attractive alternative.

REFERENCES

- Assunção, J., Rocha, R., 2019. Getting greener by going black: the effect of blacklisting municipalities on Amazon deforestation. *Environment and Development Economics*, 1–23.
- Barreto, P., Souza Jr, C., Nogueira, R., Anderson, A., Salomão, R. et al., 2006. Human pressure on the Brazilian Amazon forests. *World Resources Institute, Washington, DC*.
- Butler, R., 2008. Deforestation in the Amazon. *Mongabay.com*. <http://www.mongabay.com/brazil.html>.
- Daudt, R., Le Saux, B., Boulch, A., Gousseau, Y., 2018. Urban change detection for multispectral earth observation using convolutional neural networks. *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 2115–2118.
- De Almeida, C., Coutinho, A. and Esquerdo, J., Adami, M., Venturieri, A., Diniz, C., Dessay, N., Durieux, L., Gomes, A., 2016. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amazonica*, 46, 291–302.
- De Souza, R. A., Miziara, F., Junior, P., 2013. Spatial variation of deforestation rates in the Brazilian Amazon: A complex theater for agrarian technology, agrarian structure and governance by surveillance. *Land Use Policy*, 30, 915–924.
- De Souza, R.A., Junior, P., 2018. Improved spatial model for Amazonian deforestation: An empirical assessment and spatial bias analysis. *Ecological Modelling*, 387, 1–9.
- Deng, J. S., Wang, K., Deng, Y. H., Qi, G. J., 2008. PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing*, 29, 4823–4838.
- Dhingra, S., Kumar, D., 2019. A review of remotely sensed satellite image classification. *International Journal of Electrical & Computer Engineering (2088-8708)*, 9.

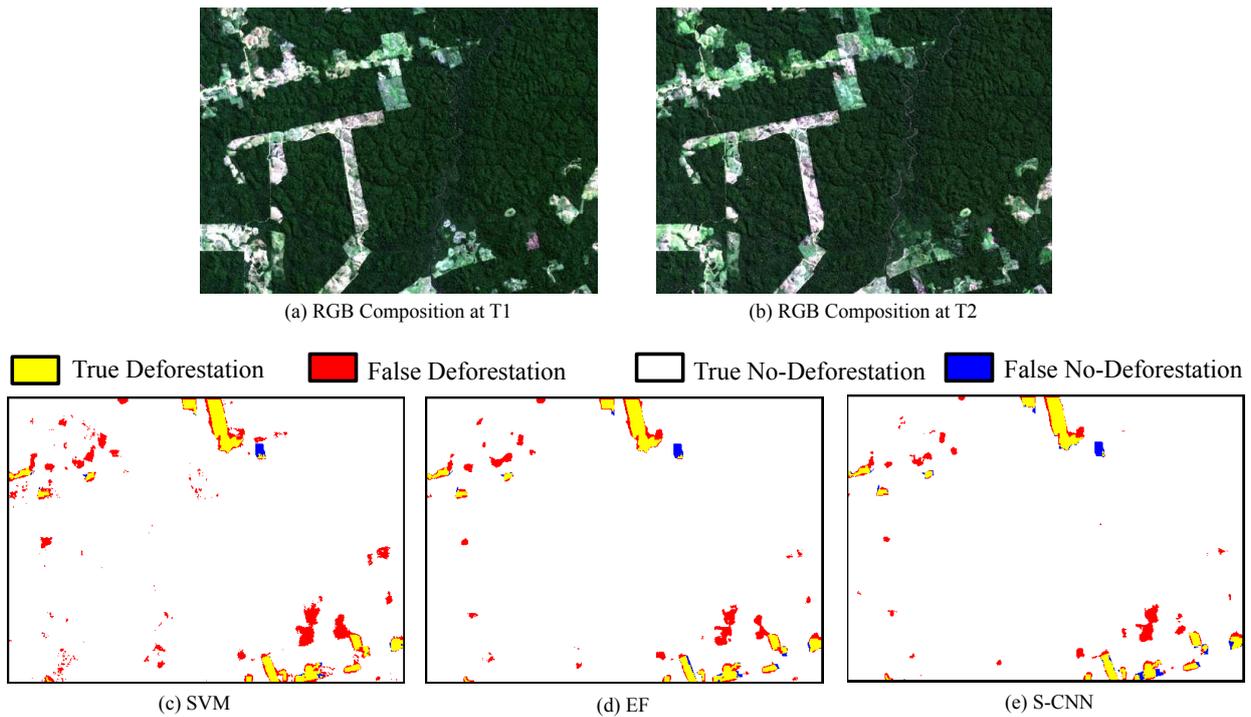


Figure 7. RGB composition and the change maps predicted by SVM, EF and S-CNN on the tile number two, which is part of the test region.

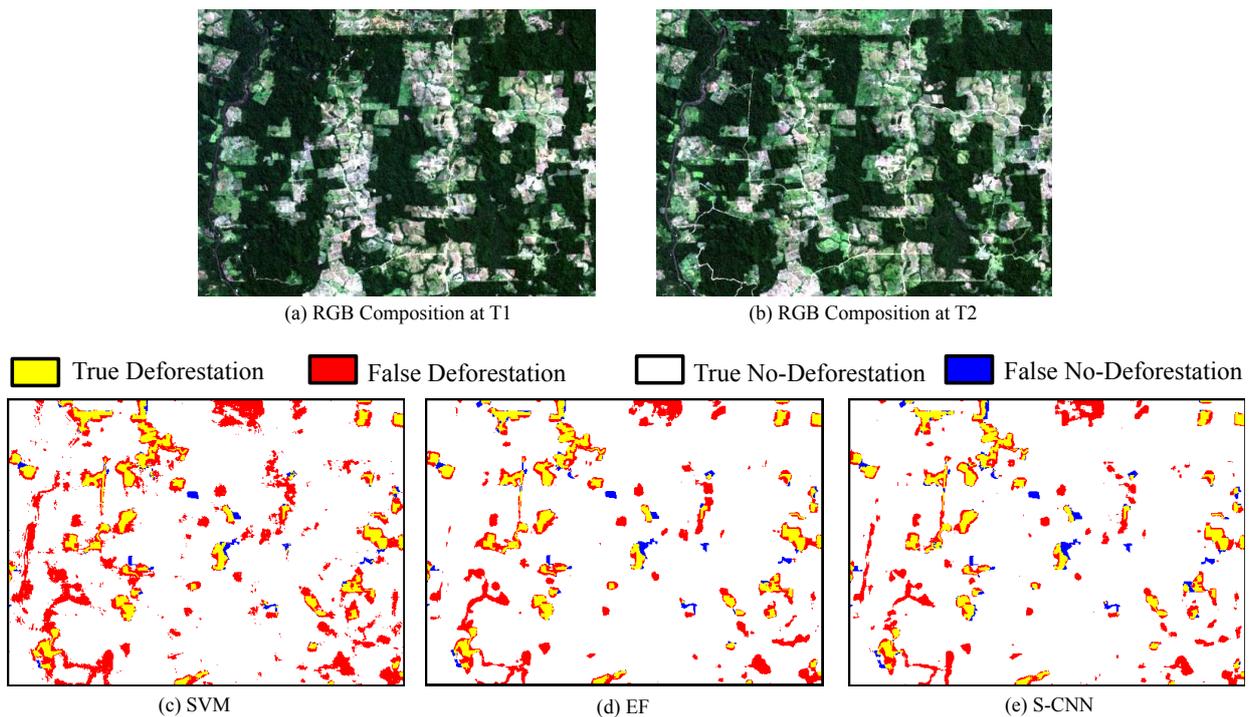


Figure 8. RGB composition and the change maps predicted by SVM, EF and S-CNN on the tile number six, which is part of the test region.

Diniz, C., de Almeida, A., Santos, D., Dias, M., da Luz, N., de Moraes, D., Maia, J., Gomes, A., da Silva, I., Valeriano, D. M. et al., 2015. DETER-B: The new Amazon near real-time deforestation detection system. *Ieee journal of selected topics in applied earth observations and remote sensing*, 8, 3619–3628.

Gola, J., Webel, J., Britz, D., Guitar, A., Staudt, T., Winter, M., Mücklich, F., 2019. Objective microstructure classification by support vector machine (SVM) using a combination of morphological parameters and textural features for low carbon steels. *Computational Materials Science*, 160, 186–196.

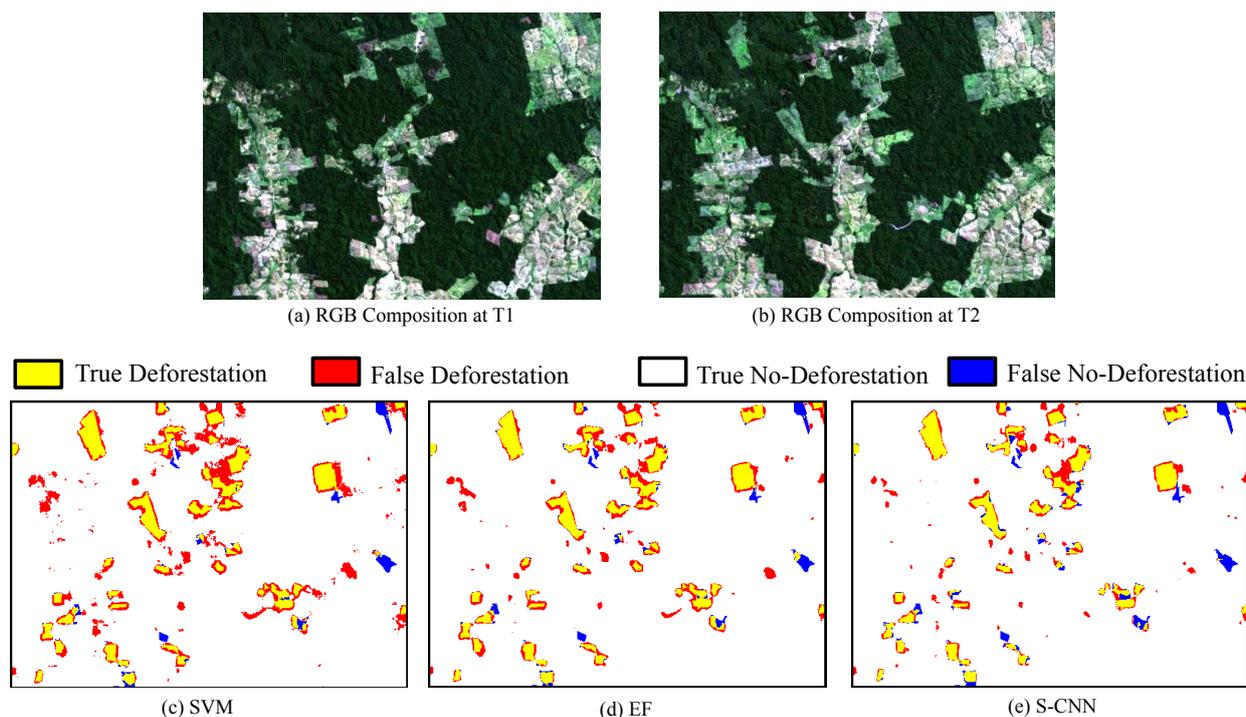


Figure 9. RGB composition and the change maps predicted by SVM, EF and S-CNN on the tile number fourteen, which is part of the test region.

Goodman, R., Aramburu, M., Gopalakrishna, T., Putz, F., Gutiérrez, N., Alvarez, J., Aguilar-Amuchastegui, N., Ellis, P., 2019. Carbon emissions and potential emissions reductions from low-intensity selective logging in southwestern Amazonia. *Forest Ecology and Management*, 439, 18–27.

Han, T., Wulder, M. A., White, J. C., Coops, N. C., Alvarez, M. F., Butson, C., 2007. An efficient protocol to process Landsat images for change detection with tasselled cap transformation. *IEEE Geoscience and Remote Sensing Letters*, 4, 147–151.

Howarth, P. J., Wickware, G. M., 1981. Procedures for change detection using Landsat digital data. *International Journal of Remote Sensing*, 2, 277–291.

Jensen, J. R., Toll, D. L., 1982. Detecting residential land-use development at the urban fringe.

Kranjčić, N., Medak, D., Župan, R., Rezo, M., 2019. Support Vector Machine Accuracy Assessment for Extracting Green Urban Areas in Towns. *Remote Sensing*, 11, 655.

Ludeke, A., Maggio, R. C., Reid, L.M., 1990. An analysis of anthropogenic deforestation using logistic regression and GIS. *Journal of Environmental Management*, 31, 247–259.

Malingreau, JP, Eva, HD, De Miranda, EE, 2012. Brazilian Amazon: a significant five year drop in deforestation rates but figures are on the rise again. *Ambio*, 41, 309–314.

Maxwell, A. E., Warner, T. A., Fang, F., 2018. Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39, 2784–2817.

Nackaerts, K., Vaesen, K., Muys, B., Coppin, P., 2005. Comparative performance of a modified change vector analysis

in forest change detection. *International Journal of Remote Sensing*, 26, 839–852.

Pal, M., 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26, 217–222.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Sathler, D., Adamo, S., Lima, E., 2018. Deforestation and local sustainable development in Brazilian Legal Amazonia: an exploratory analysis. *Ecology and Society*, 23.

Shimabukuro, Y., Duarte, V., Anderson, L., Valeriano, D., Arai, E., Freitas, R., Rudorff, B. F., Moreira, M., 2007. Near real time detection of deforestation in the Brazilian Amazon using MODIS imagery. *Ambiente e Agua-An Interdisciplinary Journal of Applied Science*, 1, 37–47.

Shimabukuro, Y. E., Beuchle, R., Grecchi, R. C., Achard, F., Miettinen, J., Simonetti, D., Gomez, M., Duarte, V., Arai, E., Anderson, L. O., 2015. Detection of forest degradation caused by fires in Amazonia from time-series of MODIS fraction images. *Curitiba: Anais XVII Simpósio Brasileiro de Sensoriamento Remoto (SBSR), INPE*.

Valeriano, D. M., Mello, E. M. K., Moreira, J. C., Shimabukuro, Y. E., Duarte, V., Souza, I. M., dos Santos, J. R., Barbosa, C. C. F., de Souza, R. C. M., 2004. Monitoring tropical forest from space: the PRODES digital project. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 35, 272–274.

Zagoruyko, S., Komodakis, N., 2015. Learning to compare image patches via convolutional neural networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4353–4361.

Zhan, Y., Fu, K., Yan, M., Sun, X., Wang, H., Qiu, X., 2017. Change detection based on deep siamese convolutional network for optical aerial images. *IEEE Geoscience and Remote Sensing Letters*, 14, 1845–1849.

Zhang, Z., Vosselman, G., Gerke, M., Tuia, D., Yang, M. Y., 2018. Change Detection between Multimodal Remote Sensing Data Using Siamese CNN. *arXiv preprint arXiv:1807.09562*.